The Journal of the Economics of Ageing 26 (2023) 100471

Contents lists available at ScienceDirect

The Journal of the Economics of Ageing

journal homepage: www.elsevier.com/locate/jeoa

Full length article

The replacement rate that maintains income satisfaction through retirement: The question of income-dependence

Julian Schmied *

Free University of Berlin, Kaiserswerther Str. 16-18, 14195 Berlin, Germany
Max Planck Institute for Demographic Research, Konrad-Zuse Str. 1, 18056 Rostock, Germany

A R T I C L E I N F O

JEL classification:
J26
C14
H55
J32
D19

Keywords:
Income satisfaction
Saving adequacy
Retirement
Pensions
Inequality

A B S T R A C T

Benchmark replacement rates are commonly used to set up saving plans or to assess retirement preparedness. An open question is whether high earners need the same replacement rate as low earners. In this paper, I apply the GAESE framework, an approach known from the equivalence scale literature, to assess how the replacement rate that maintains income satisfaction through retirement relates to income levels. Using longitudinal data from the German Socio-Economic Panel, and applying fixed-effects ordered logit models, I find that the benchmark replacement rate decreases with income. For singles, this finding is consistent across many modifications of the approach, whereas for couples the finding is sensitive to the composition of the retiree household, i.e. whether or not the retiree's partner is still employed.

Introduction

In many aging countries, private savings have become an essential pillar in the retirement income portfolio. The shift from public to private decision-making is accompanied by a lot of uncertainty, and requires guidance. The replacement rate – i.e., the percentage of the end-of-career employment income that is replaced by the retirement income – is a key parameter for pension planning tools based on life-cycle models, in which it represents the decline in income that the individual or the household is willing to accept after retirement (Skinner, 2007; Scholz et al., 2006). But what is a good choice for this parameter?

In practice, 70% of net income (Schulz and Carrin, 1972) is often used as a benchmark replacement rate. Dudel et al. (2016) estimated for Germany that 86% of net income is needed. Both approaches assume that one benchmark fits all income levels. The survey literature, on the other hand, suggests that people with different income levels need different replacement rates (Binswanger and Schunk, 2012).

Benchmark replacement rates are also important in public policy. For example, the UK Pension Commission 2004 recommends replacement rates of 80% for annual incomes below PBS 9500, and 50% for incomes above PBS 40,000. There is no empirical justification for these values except that those are the realized replacement rates of the recent cohorts of retirees. Arguably, these figures do not reflect whether or not the respective households are satisfied with their financial situation after retirement (Crawford and O'Dea, 2012).

In this paper, I estimate a benchmark replacement rate that has an empirical foundation, and that varies with income. I follow Dudel et al. (2016), and estimate a replacement rate that maintains income satisfaction through retirement, but I use a more flexible specification that allows the rate to vary with income. The approach is based on the Generalized Absolute Equivalence Scales Exactness framework (GAESE), which was originally applied to derive income-dependent scales that equalize income across families of different compositions (Donaldson and Pendakur, 2006). The common key assumption is that welfare can be directly observed in the data.

The method is applied as follows: I track a sample of people who are 50 years or older over time. Second, I define a number of household situations that vary in terms of retirement status and household size. Third, I measure the relationship between household income and income satisfaction within the respective household types. I also pick a reference household; e.g., for single households, employed single individuals. Finally, I estimate a shift and a scaling parameter by Maximum Likelihood that can, in turn, be used to plot the retirement income ratio against income.

As self-reported income satisfaction is the dependent variable of the regression analysis, I apply longitudinal ordered response models,

* Correspondence to: Max Planck Institute for Demographic Research, Konrad-Zuse Str. 1, 18056 Rostock, Germany.
E-mail address: schmied@demogr.mpg.de.

https://doi.org/10.1016/j.jeoa.2023.100471

Available online 7 August 2023

as discussed in Baetschmann et al. (2015), that take into account both the ordinal nature of the dependent variable, and the individual unobserved but time-invariant heterogeneity. The approach does not make strong assumptions regarding the interpersonal comparability of income evaluations, but instead only assumes that individuals are consistent over time. Further, as utility is measured on the individual level, no strong assumptions about intra-family resource allocation are needed (Chiappori, 2016).

I investigate Germany, which has a pay-as-you-go pension system of the Bismarckian variety in which benefits strongly depend on contributions, and retirement incomes are still mainly provided by statutory pensions. I use longitudinal data from the German Socio-Economic Panel (SOEP), which has collected yearly data on subjective income evaluations for nearly 30 years. The sample consists of 114,756 observation-years.

I find that, among singles, the replacement rate that maintains income satisfaction through retirement falls, within the interquartile range of earnings, from 67% for monthly household net earnings of around EUR 1300 to about 59% for earnings around EUR 2500. For couple households, the metric falls from 80% for joint earnings of around EUR 2800 to 72% for earnings around EUR 5100. The income-independence assumption is rejected on the 1%-Level for both single-person households and couple households. Results for couple households, however, depend on the labor market status of respondents' household partner.

In the sensitivity analysis, I tested transformations of the dependent variable, additional time-varying covariates, two alternative identification strategies, stratifications along time-invarying variables, and modifications of the age threshold. The results for singles turned out to be fairly consistent across the models.

Overall, the results suggest that households at the bottom level of the income distribution have higher fixed costs to replace when they retire. They also confirm the findings of Binswanger and Schunk (2012) using qualitative data from the US and the Netherlands. From a policy perspective, the results call into question whether constant benefits-to-payments ratios, which has long been applied in Germany, is an approach in which the majority of the population is financially satisfied after retiring. Households with lower incomes have to save more proportional to their earnings, and because these households usually have fewer opportunities for wealth accumulation (Bernheim et al., 2001), they have to be supported in their efforts to save.

By using a model that enables me to examine whether replacement rates depend on income levels I add to the body of existing literature. The previous studies that were closest to this one were conducted by Dudel and Schmied (2019) and Binswanger and Schunk (2012). In the former study, the authors investigated how much the replacement rate needs to be to maintain consumption levels. They also tested whether it is fair to use one benchmark for all income levels. They applied econometric tests which indicate whether income-independence must be rejected, but not whether replacement rates increase or decrease with income. The authors found mixed results depending on the allowed flexibility of the test. In a different study design, Binswanger and Schunk (2012) asked a sample of pre-retirees from the US and the Netherlands how much money they would need to maintain an adequate standard of living in retirement, given their current income. In both countries, low-income households expressed a desire for a larger fraction of their income, suggesting a decreasing gradient in income. While this study design has many benefits and relies on weak assumptions it is still an ex-ante approach. The question whether individuals desire a lower or a higher replacement rate *after* they retire remains unresolved, as they may change their minds. Thus, an ex-post analysis like to one conducted in this paper is an important complementary study design.

The remainder of this paper is structured as follows. In Section "Related literature", I review the current literature on replacement rates and corresponding benchmarks. In Section "Methodology", I describe the econometric framework and the identification strategy I employ. In Section "Data", I describe the dataset I use. I present the main findings in Section "Result"s and offer a wide range of robustness tests. In Section "Discussion", I discuss the results and make some attempts to explain them. Section "Conclusion" concludes.

Related literature

Replacement rates, sometimes referred to as retirement income ratios, are used in various ways and for different purposes. For individuals, replacement rates may be used to provide a projection of their future living standards, as the rates place their projected income after retirement in relation to their income while working. The average replacement rate of a population, such as the population of a country, may be used to assess savings adequacy across cohorts (Smith, 2003; Geyer and Steiner, 2014; Knoef et al., 2016). Similarly, replacement rates are often used to study the effects of policy reforms on retirement incomes (Palmer, 1989), or as indicators of pension adequacy when comparing countries with different pension systems (OECD, 2015). In practice, actuaries and financial advisers, as well as online retirement planners, use replacement rates as benchmarks to set up individual saving plans, or to assess the adequacy of current accumulated wealth (Skinner, 2007). At both the individual level and the population level, the assessment of realized replacement rates against a benchmark replacement rate may be used as a measure of economic well-being in old-age (Dudel and Schmied, 2019). While realized replacement rates can be observed in register data and are usually higher at lower income levels, the question of whether benchmark replacement rates follow the same pattern remains open.

In the first report of the UK Pension Commission (2004), the authors suggested the use of a benchmark replacement rate that decreases with income. They recommended a threshold of 80% for annual incomes below PBS 9500, and a threshold of 50% for incomes above PBS 40,000. The commission justified the choice of these thresholds by stating that they are the actual replacement rates of current cohorts. However, it is not clear why the realized replacement rates were selected as the benchmark replacement rates, given that the observed households might have failed to meet their financial retirement goals, or could be unsatisfied with their current resources.

Dudel et al. (2016) established a benchmark replacement rate based on an explicit objective, the maintenance of income satisfaction through retirement. Using German SOEP data from 1989–2014, they estimated this rate to be around 90%, and assumed that it applies equally to people with low and high incomes. Dudel and Schmied (2019) tested this assumption with cross-sectional expenditure data and found no clear evidence of income-dependence.

Binswanger and Schunk (2012) uncovered a different pattern using a customized questionnaire, i.e., they asked employed individuals how much of their current income (or a projection thereof) they would need to maintain an adequate standard of living in retirement. In both examined countries, the US and the Netherlands, low-income households expressed a desire for a larger fraction of their income, which suggests a decreasing gradient in income. Interestingly, the ranges in the two countries were very different. In the US, people in the lowest quintile expressed a preference for a rate of around 108%, whereas people in the top quintile expressed a preference for a rate of around 54%. In the Netherlands, the range was much narrower, from 69% to 63%.

Similar inconsistencies between empirical and qualitative evidence have been observed in the equivalence scale literature. Equivalence scales are important metrics for standardizing income between households of different compositions (Deaton and Muellbauer, 1980). While income-independent scales had been popular for a long period of time, qualitative evidence has since indicated that the income-independence assumption must be rejected (Koulovatianos et al., 2005).[1] Therefore, Donaldson and Pendakur (2006) later proposed a more general framework that allowed equivalence scales to vary with income. At that point, equivalence scales were mostly based on revealed preferences from expenditure systems, and were estimated using detailed expenditure data. In more recent work, Biewen and Juhasz (2017) adjusted the approach so that it is applicable to satisfaction data. Many of the methodological questions discussed in the equivalence scale literature can be also applied to efforts to estimate the replacement rates that maintain income satisfaction (Dudel et al., 2016).

Methodology

Essentially, the main idea of the approach is as follows: I track each person's income and level of satisfaction with their income in the years leading up to retirement, I then accompany the person into retirement, where they once more express a degree of satisfaction. Consider, for example, a scenario where a person's level of satisfaction remains unchanged but his/her income declines by 30%. In that situation, the approach suggests that the person needs a replacement rate of 70% because s/he needs 30% less income to feel the same level of income satisfaction. This paper explores whether a wealthier person would also require 70% or less or more.

More generally, the identification is based on data from individuals living in different household situation, where one situation is while working while the other is in retirement. The replacement rate averages over the adjustments the individuals make during the transition while utility is set equal.

The main goal of the paper is to check whether or not this replacement rate is better represented by an income-dependent specification. I use the framework by Biewen and Juhasz (2017) where different functional forms of equivalence scales, including income-dependent and income-independent forms, can be directly compared.[2]

Let u_{it} denote satisfaction with the household income y_{it} of individual i in year t. u_{it} is determined in this model by equivalent income eq_{it}; a vector of other time-varying covariates z'_{it} such as age and health, individual-specific time-invariant effects such as whether a person is generally optimistic ϕ_i; and an error term ϵ_{it} that captures measurement error and temporary shocks to u.[3,4]

$$u_{it} = \beta_1 log(eq_{it}) + z'_{it}\gamma_k + \phi_i + \epsilon_{it} \tag{1}$$

Let w^k denote a vector of $k+1$ binary variables $I()$ that indicate whether an individual i is observed in a specific household situation in year t. The household situation is determined by the number of individuals living in the same household and their retirement status. In the reference household situation w^0, the individual or couple is still employed (see Section "Household situations"). The comparison situations $w_1, w_2, \ldots, w_k$ correspond to retirement household situations. It is important to note that when considering couples, apart from the situation where both partners are retired, other scenarios can also be interpreted as a *retired couple*. For instance, the respondent could be retired while the partner is unemployed.

$$eq_{it} = f(w_{it}, y_{it}) \tag{2}$$

[1] The assumption was referred to as base independence (IB; Lewbel, 1989b) or equivalence scale exactness (ESE; Blackorby and Donaldson, 1993).

[2] Conceptually, equivalence scales are equal to replacement rates that maintain utility. However, the latter implies constant household sizes, where the transition does not manifest itself in different family sizes but in different retirement statuses.

[3] For example, unpleasant weather on the day of the survey might affect the general mood of the respondents (see Kahneman and Krueger, 2006, for more examples).

[4] The functional relationship of the equivalent income is specified as the logarithmic income and linear income satisfaction. That approach provides the best fit for the data (see Appendix A).

The equivalent income eq_{it} denotes the retirement income that is needed to restore the income satisfaction u from the reference situation. $f()$ is a transformation function showing how individuals adjust their income y_{it} when moving from the reference to other comparison situations.

In the reference situation, equivalent income is simply the observed household income:

$$f(w^0_{it}, y_{it}) = y_{it} \tag{3}$$

The replacement rate that maintains the income satisfaction through retirement is defined by the ratio of the income in the reference situation and equivalent income:

$$r(w_{it}, y_{it}) = \frac{y_{it}}{eq_{it}} \tag{4}$$

From (3) and (4) follows that the replacement rate of the reference situation is 100%:

$$r(w^0_{it}, y_{it}) = 1 \tag{5}$$

While this is true for any form of $f()$, there are many different transformation functions (2). In an income-independent setting (IB)[5] equivalent income may be obtained by

$$eq_{it} = \frac{y_{it}}{\rho(w_{it})}, \tag{6}$$

where

$$\rho(w_{it}) = 1 \times I(w_0 = 1) + b_1 \times I(w_1 = 1) + \cdots + b_k \times I(w_3 = 1). \tag{7}$$

Starting with 1 for the reference situation, replacement rates can be directly observed for all k comparison households from the estimators $b_1, b_2, \ldots, b_k$. These are constant and independent of income, as

$$r^{IB}(w_{it}) = \frac{y_{it}}{eq_{it}} = \frac{y_{it}}{\frac{y_{it}}{\rho(w_{it})}} = \rho(w_{it}). \tag{8}$$

The GAESE specification allows eq_{it} to have a fixed component $\alpha(w_{it})$ and a variable component $\rho(w_{it})$ of the income difference that is needed to maintain u when moving from the reference household situation to the comparison household (Donaldson and Pendakur, 2006).

$$eq_{it}(w_{it}, y_{it}) = \rho(w_{it})y_{it} + \alpha(w_{it}) \tag{9}$$

where

$$\alpha(w_{it}) = 0 \times I(w_0 = 1) + a_1 \times I(w_1 = 1) + \cdots + a_k \times I(w_k = 1) \tag{10}$$

Here, the change of the equivalent income function eq_{it} is constant with respect to the household income y_{it} given the household type w_{it} (Biewen and Juhasz, 2017).

$$\frac{\partial eq_{it}(w_{it}, y_{it})}{\partial y_{it}} = \rho(w_{it}) \tag{11}$$

Following (4), the replacement rate from the GAESE form is obtained by

$$r^{GAESE}(w_{it}, y_{it}) = \frac{y_{it}}{eq_{it}} = \left(\rho(w_{it}) + \frac{\alpha(w_{it})}{y_{it}}\right)^{-1} \tag{12}$$

Whether or not r varies with income is indicated by the fixed component $\alpha(w_{it})$. If it is negative, the replacement rates are higher for poorer households, and vice versa. If it is zero, the metric collapses to the inverse of (6) and becomes independent of the income level. Therefore, examining the significance of α is an econometric test to check for income independence (Biewen and Juhasz, 2017).

Using the conditional likelihood estimator (CLM, Chamberlain, 1979) the coefficients of (1) can be estimated despite the non-linearity

[5] Referring to the independence of base assumption from the equivalence scale literature, see Lewbel (1989a).

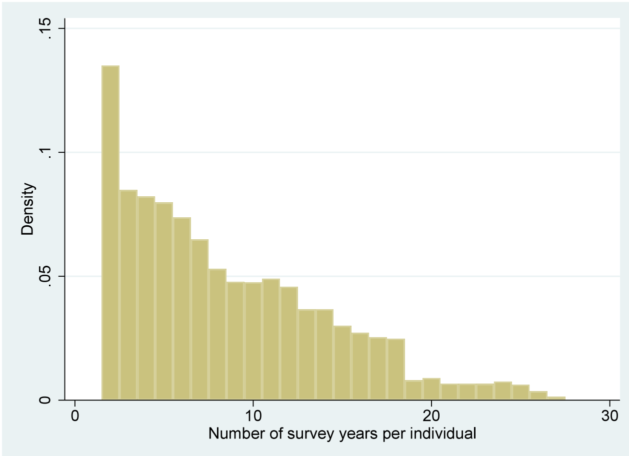

Fig. 1. Number of observed years per individual.
Source: SOEP 1991–2017.

of eq_{it}.[6] The CLM is based on a binary logit model which takes into account individual time-invariant effects. With the blow-up-and-cluster (BUC) method, the CLM can be effectively applied to ordinal structured dependent variables (Baetschmann et al., 2015). In Appendix C I provide details on the implementation of the BUC.

The estimation strategy makes the following assumptions: utility can be directly observed in the data, here by income satisfaction; income satisfaction can be compared within individuals (discussed in the next section); the true utility replacement rate is non-linear (tested in Section "Main results"); and r is sufficiently controlled for by observables Z_{it} or time-invariant unobserved heterogeneity α_i (discussed in Section "Robustness").[7]

Data

Sample

I use data from the German Socio-Economic Panel (SOEP), which includes over 15,000 households and 30,000 individuals in the 34th wave. I use observations until the survey year of 2017. Individuals are followed over time and not across households. Importantly, this implies that for every household with more than one adult, each adult is surveyed independently.

I reduce the sample to household situations in which the respondent is aged 50 or older, based on the fact that replacement rates usually refer to the individuals' earnings in the final years of their career. This threshold is somewhat arbitrary, and I discuss how other thresholds could affect the outcomes in Section "Robustness". I drop households with children living in the household, and households with more than two adults. Finally, I drop households with missing information on key variables. That leaves an unbalanced panel of 14,165 households with at least two survey years and a median of seven observed years ranging from 1991 to 2017 (see Fig. 1).

Household situations

In the next step, I construct binary variables to define single and couple household situations. Essentially, these households differ according to the number of persons living in the household and their labor force status. The household situations are distinct from each other; i.e., one household cannot be in two or more household situations at the same time. For every couple household, there are two individual observations, but in the analysis, the standard errors are clustered on the household level.

The reference category for singles a_{emp} denotes an employed single household situation in which the respondent is employed and did not receive any income from pensions in the previous year (7959 observation-years).[8] a_{ret} denotes a single retiree household situation in which the respondent is not employed and received pension benefits in the previous survey year (22,784 observation-years).

The reference category for couples aa_{emp} denotes a double-earner household situation in which both the respondent and his/her partner are employed, and neither received no pension benefits in the previous year (18,691 observation-years). aa_{ret} denotes a two-retiree household situation in which neither the respondent nor his/her partner is employed and both received pension benefits in the previous year (40,691 observation-years). $aa_{ret;emp}$ denotes a one-earner–one-retired household situation in which the respondent/partner is employed and received no pension benefits in the previous year, while the partner/respondent is retired and non-employed (9778 observations years). $aa_{ret;un}$ denotes a one-retired–one-unemployed household situation in which both the respondent and his/her partner are non-employed, and only one of them has received pension benefits in the previous year (7338 observation-years).

Additional household situations that fall within the defined age range and do not include retired people are used as control variables in the model but are not given any interpretation. Specifically, single or couple households in which all respondents are unemployed (a_{un}, aa_{un}) as well as one-earner–one-unemployed household situation in which the respondent/partner is employed while the partner/respondent is unemployed and did not receive any pension benefits in the previous year ($aa_{emp;un}$).

The transition frequencies within and between each defined household situation are displayed in Table 1. Among singles, there are 402 transitions from employment into retirement. Among couples, there are 327 transitions from a situation where they are both employed to one where they are both retired. Given the 50k observation-years in aa_{emp} and aa_{ret}, this may seem insignificant, but it is not surprising given the small number of couples who retire during the same survey years. More people move into households where one person is retired and the other is still working (938); less people into household where one person is retired and the other unemployed (110). Finally, Table 1 demonstrates that the number of households that change their membership is relatively small (see first row).

The average number of survey years among singles is 4.7 prior to the retirement transition and 6.8 following it. The average observation period for couples is 4.6 survey years for those who are both employed, 7.1 for those who are both retired, 3.5 for those who have one in retirement and the other in employment, and 3.7 years for those who have one retired and the other unemployed.[9]

[6] Biewen and Juhasz (2017) provide Stata code for implementation.

[7] One additional key assumption for equivalence scales that are based on families or multiple-person households is that there is a single utility function for the household/family as a whole even though it is well known that resources are not equally shared among household members (Chiappori, 2016). In this application that assumption is not needed, as in the SOEP data, the utility of all (adult) household members is individually assessed (households with children are not part of the sample).

[8] Dudel et al. (2016) showed that using alternative concepts of retirement definition leads practically to the same results.

[9] This is based on whether people are actually relocating the respective states. In other words, if there have been no survey years in the relevant states, they are not included. These significantly reduce the averages when included.

Table 1
Transition frequencies between and within household situations.
Source: SOEP 1991–2017.

	a_{emp}	a_{ret}	a_{un}	aa_{emp}	aa_{ret}	$aa_{ret;emp}$	$aa_{emp;un}$	$aa_{ret;un}$	aa_{un}
a_{emp}	6609	402	297	94	11	18	15	6	3
a_{ret}	16	22395	115	0	61	21	0	9	0
a_{un}	192	328	1089	2	3	3	4	1	0
aa_{emp}	104	18	7	14115	327	938	974	110	39
aa_{ret}	0	845	3	0	35189	10	0	119	2
$aa_{ret;emp}$	55	57	6	12	1201	6886	62	530	4
$aa_{emp;un}$	33	7	19	541	156	470	3677	393	214
$aa_{ret;un}$	3	97	40	2	1241	229	5	5941	14
aa_{un}	2	3	8	4	60	6	63	186	352

Dependent variable

This approach is based on the assumption that utility is directly observed in the data. To assess the utility of all members of the household, self-reported evaluations of the household's financial situation are used. The SOEP has included such a question since 1989, and thus, almost from the beginning the of the survey. Respondents are asked how satisfied they are with their household income on a scale from zero to 10, with 10 being the most satisfied. Note again that for a couple household there are two different evaluations of the same household income.

The use of subjective evaluations of income has been a popular approach for conducting welfare analysis in general, and for examining equivalence scales and pension adequacy in particular (e.g., Ferrer-i Carbonell and Frijters, 2004; Pradhan and Ravallion, 2000; van Praag and Kapteyn, 1973; Schwarze, 2003; Biewen and Juhasz, 2017; Koulovatianos et al., 2005; Binswanger and Schunk, 2012). While the literature has been less critical of self-reported evaluation of income (or other domains) in the recent years, there are a number of assumptions and potential problems that should be taken into account when using this approach (e.g., Krueger and Schkade, 2008; Bertrand and Mullainathan, 2001; Layard et al., 2008). As I will discuss in the following, most of these reservations can be addressed within the applied study design.

First, misreporting can be systematically correlated with unobserved or observed individual characteristics (Bertrand and Mullainathan, 2001). For example, it may be the case that individuals with higher financial literacy have systematically different assessments of their financial situation than individuals with lower financial literacy. This problem is addressed in this paper by the fixed-effects estimation, which allows for correlations of measurement error with time-invariant characteristics. Problems of measurement errors are accelerated if both the dependent and independent variable are subjective (Ferrer-i Carbonell and Frijters, 2004). Here I use household income which is sufficiently objective and thus this is not a concern in this application.

Second, a number of studies have shown that the design of the questionnaire can affect the outcome variable (see Kahneman and Krueger, 2006, for a literature review). Most importantly, there is evidence that the ordering, the context, and the vagueness of the question can change the results (Bertrand and Mullainathan, 2001). It appears that using a specific domain (household income), rather than general life satisfaction, is less problematic (Krueger and Schkade, 2008). Moreover, it can be expected that respondents beyond retirement age understand the question as it is posed in a straightforward manner (see above). The question is also asked at the very beginning of the questionnaire, and is, unlike in other aging datasets, included in the core study. Finally, the question relates to each household's current financial situation, and not to their past or future situation (Pudney, 2011).

Next, there is the issue of how the variable should be treated in the econometric model, and what assumptions should accompany it. This issue is discussed extensively in Dudel et al. (2016) and in Section "Methodology". Essentially, the assumption of cardinality implies that an increase in satisfaction of one point is the same, regardless of whether the increase is from two to three or from four to five. In this paper, I apply different types of methods that assume either cardinality or weak ordinality.

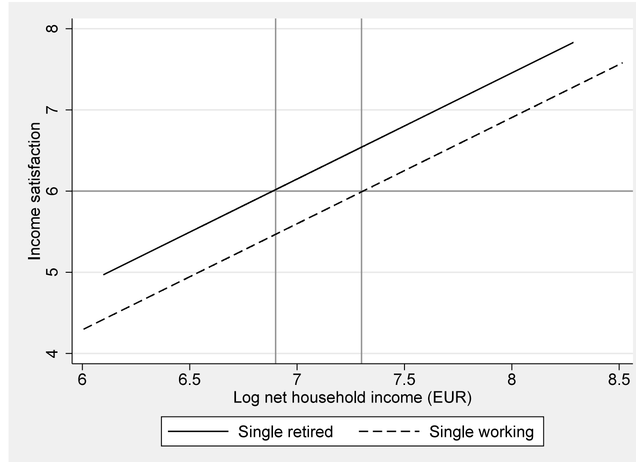

Fig. 2. Linear Engel curves and base dependence.
Note: The lines represent linear predictions of a fixed effect regression of income satisfaction on the log of household income.
Source: SOEP 1991–2017.

Income and other explanatory variables

Household income is a key variable in the analysis. My income variable is based on what the SOEP calls household post-government income, which is reported by either the respondent or the household head. As I examine a long time period, I deflate income with the consumer price index provided by the Federal Statistical Office. Using after-tax income is key to estimating a convincing replacement rate, because the way income is taxed after retirement has changed in Germany. To make the replacement rate as close as possible to the actual income that is generated by the household, all sources of income should be taken into account. In addition to labor earnings, asset flows, private retirement income, private transfers, and social security pensions are used. Labor earnings include wages and salary from all employment, including training income, self-employment income, bonuses, overtime, and profit-sharing income. Asset flows include income from interest, dividends, and rent. Private transfers include payments from individuals outside of the household including alimony and child support payments. Social security pensions include payments from old age, disability, and widowhood pension schemes (Grabka, 2020). The variable does not take imputed rent into account. Homeownership is discussed in Section "Robustness".

Extreme values of household income are excluded by trimming the top and the bottom 1% of the income distribution. I do that for every household situation separately.

In the baseline model, I control for age and period dummies. Note that the applied models do not allow for the inclusion of time-invarying covariates, such as gender, cohort, or education. I examine the sensitivity of the results by including time-varying variables that affect income satisfaction around the time of retirement, such as health in Section "Robustness".

Results

Graphical analysis

I start the analysis with a graphical analysis that is known from the literature on Engel curves (e.g. Deaton and Muellbauer, 1986). It serves

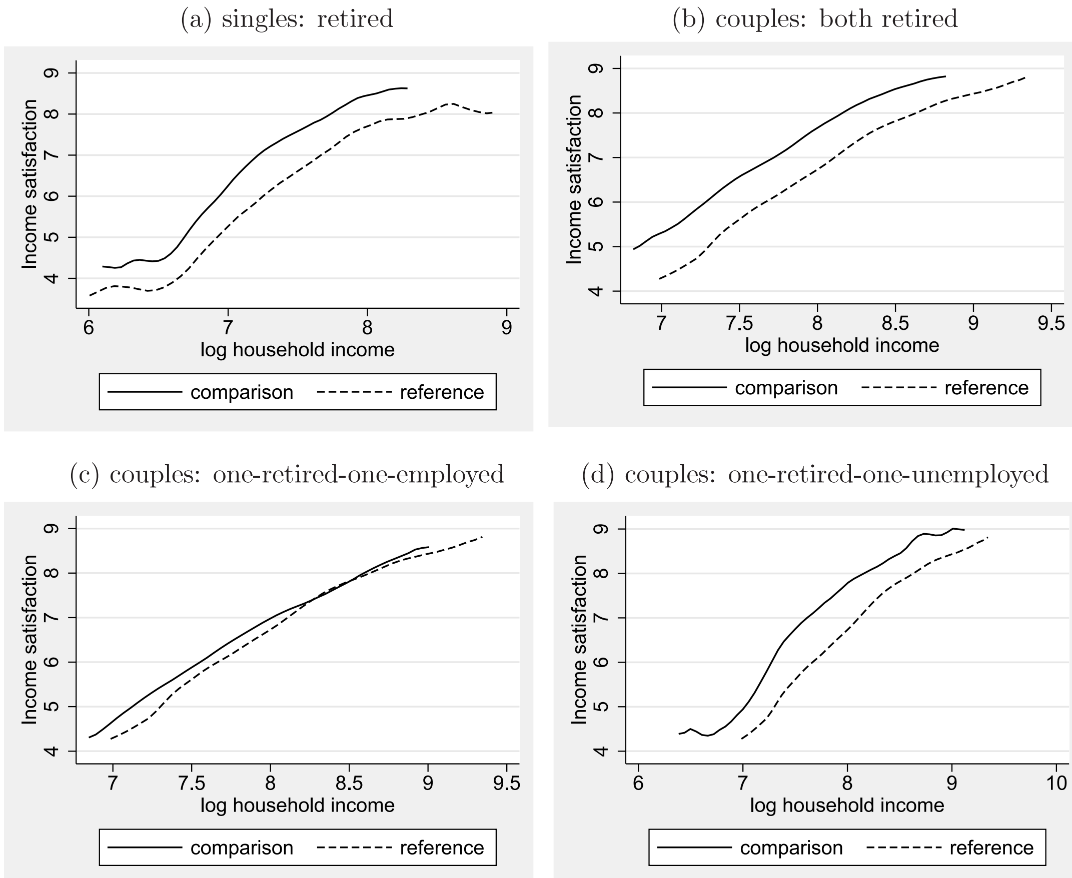

Fig. 3. Non-parametric regression lines of household income and income satisfaction by household type.
Note: The reference household for panel (a) is a single individual in employment; for panel (b) to (c) the reference household is a household in which both are in employment.
Source: SOEP 1991–2017.

as a non-parametric test for the replacement rates' independence from income and aims to aid the reader in understanding the identification strategy.

Consider Fig. 2 in which two linear regression curves are used to compare the relationship between the log of income and income satisfaction. For simplicity, I just consider singles and disregard any time-varying affects such as age.[10] The income difference that causes the regression line of the reference household (solid line) to shift horizontally until it reaches the comparison household (dashed line) defines the (log of the) replacement rate. It is below 100% if the regression line of the comparison household is located on the left-hand side of the regression line of the reference household.

For example, an individual that has – at the end-of-the career – an income satisfaction level of six (horizontal gray line) and earnings of $exp(7.3) = 1,480$ would need a retirement income of $exp(6.9) = 992$ to maintain the satisfaction level through retirement. The replacement rate that maintains that satisfaction level in that case is $992/1,480 = 67\%$. In this setting, the rate applies to all income levels, but this type of identification is only valid if the regression curves are linear.

Fig. 3 shows non-parametric fits of the relationship between income and income satisfaction for both single and couple household situations. Panel (a) casts doubt on the claim that a linear fit as used in Fig. 2 is an accurate representation of the data. Both regression lines show a decline for low incomes and an additional rate of change at the median income, which in the sample is at exp(7.2).

Fig. 4 shows the result of estimating the GAESE specification (12) without control variables, i.e., $\hat{u_{it}} = \hat{\beta} \times log(\hat{\rho} \times y_{it} + \hat{\alpha}) + \hat{\epsilon}_i$ for retired

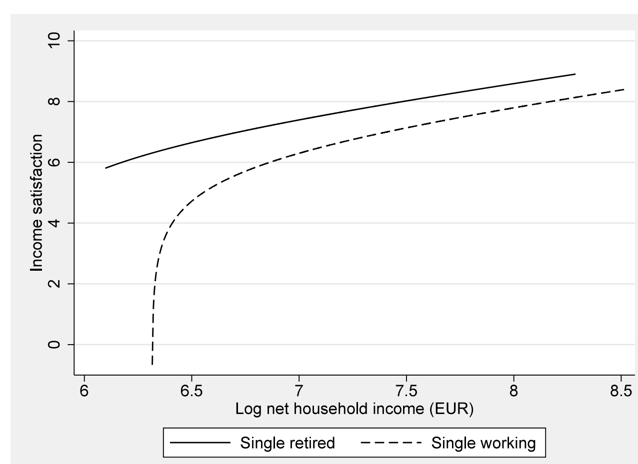

Fig. 4. Linear Engel curves and base dependence.
Note: The lines represent linear predictions of a fixed effect regression of income satisfaction on the log of household income.
Source: SOEP 1991–2017.

singles and working single separately. Note that α determines how curvy the regression lines are, so it is easy to see that when α is zero, the lines become linear and replacement rates income independent.

Contrasting Figs. 4 and 3 Panel (a) shows that while the rate of change at low incomes is not captured, the predicted values of the GAESE specification nicely reflect the declining marginal rate of

[10] Income satisfaction is assumed to be cardinal in this analysis.

Table 2
Coefficients Model (12) ; Single worker as reference category.
Source: Socio-Economic Panel 1991–2017.

Parameter	Variable	Coefficient	SE (clustered)	P>z
β_1	log(Income)	1.010	0.069	<0.001
$\alpha^{a_{ret}}$	scaling comp retired singles	−486.8	127.4	<0.001
$\alpha^{a_{un}}$	scaling comp unemployed singles	132.1	48.7	0.007
$\alpha^{aa_{ret}}$	scaling comp both retired	−1203.6	349.2	0.001
$\alpha^{aa_{emp}}$	scaling comp both working	−1519	458.1	0.001
$\alpha^{aa_{ret;emp}}$	scaling comp retired/working	−768.6	305.0	0.001
$\alpha^{aa_{emp;un}}$	scaling comp working/unemployed	−544.8	132.1	0.012
$\alpha^{aa_{ret;un}}$	scaling comp retired/unemployed	−416.8	132.4	0.002
$\alpha^{aa_{un}}$	scaling comp both unemployed	95.2	72.3	<0.188
$\rho^{a_{ret}}$	translating comp retired singles	1.881	0.201	<0.001
$\rho^{a_{un}}$	translating comp unemployed singles	0.388	0.066	<0.001
$\rho^{aa_{ret}}$	translating comp both retired	1.844	0.274	<0.001
$\rho^{aa_{emp}}$	translating comp both working	1.626	0.252	<0.001
$\rho^{aa_{ret;emp}}$	translating comp retired/working	1.414	0.209	<0.001
$\rho^{aa_{emp;un}}$	translating comp working/unemployed	0.936	0.115	<0.001
$\rho^{aa_{ret;un}}$	translating comp retired/unemployed	1.085	0.138	<0.001
$\rho^{aa_{un}}$	translating comp both unemployed	0.300	0.076	<0.001
γ_1	Age	−0.017	0.005	<0.001

Note: The dependent variable is income satisfaction scaled $0-10$. Period effects are included but not shown. Maximum likelihood estimation. Number of cluster: 8,436. Standard errors are clustered on the household levels.

Table 3
Coefficients Model (12) ; Dual-earner couple as reference category.
Source: Socio-Economic Panel 1991–2017.

Parameter	Variable	Coefficient	SE (clustered)	P>z
β_1	log(Income)	1.117	0.083	<0.001
$\alpha^{a_{ret}}$	scaling comp retired singles	−386.6	122.1	<0.002
$\alpha^{a_{un}}$	scaling comp unemployed singles	151.5	56.7	<0.008
$\alpha^{a_{emp}}$	scaling comp employed singles	−198.6	82.7	<0.016
$\alpha^{aa_{ret}}$	scaling comp both retired	−871.6	287.0	0.002
$\alpha^{aa_{ret;emp}}$	scaling comp retired/working	−360.9	263.6	0.171
$\alpha^{aa_{emp;un}}$	scaling comp working/unemployed	−370.5	135.4	0.006
$\alpha^{aa_{ret;un}}$	scaling comp retired/unemployed	−296.1	138.6	0.033
$\alpha^{aa_{un}}$	scaling comp both unemployed	122.3	82.2	<0.137
$\rho^{a_{ret}}$	translating comp retired singles	1.746	0.167	<0.001
$\rho^{a_{un}}$	translating comp unemployed singles	0.404	0.076	<0.001
$\rho^{a_{emp}}$	translating employed singles	1.129	0.114	<0.001
$\rho^{aa_{ret}}$	translating comp both retired	1.577	0.162	<0.001
$\rho^{aa_{ret;emp}}$	translating comp retired/working	1.148	0.118	<0.001
$\rho^{aa_{emp;un}}$	translating comp working/unemployed	0.843	0.073	<0.001
$\rho^{aa_{ret;un}}$	translating comp retired/unemployed	0.989	0.101	<0.001
$\rho^{aa_{un}}$	translating comp both unemployed	0.326	0.079	<0.001
γ_1	Age	−0.018	0.067	<0.001

Note: The dependent variable is income satisfaction scaled $0-10$. Period effects are included but not shown. Maximum likelihood estimation. Number of cluster: 8,436. Standard errors are clustered on the household levels.

income in income satisfaction for household incomes above EUR 1400. Goodness-of-fit measures support this assertion (see Section "Goodness-of-fit").[11]

For couples, the graphical analysis is not as straightforward as for singles. Panels (b) to (d) of Fig. 3 show non-parametrically fitted values for couples with one or more retirees. Although, linearity cannot be completely ruled out by visual inspection alone, the regression lines for the reference households in which both household members are employed (dashed lines), as well as the comparison households from Panel (b) and (c) show a nuance decline. Panel (d) looks similar to Panel (a) and is curvier, but this could be because there weren't many of that kind of household in the sample.

Main results

Tables 2 and 3 display the regression results for Eq. (12), with a household with one worker as the reference category and a household with two workers as the reference category, respectively. The model estimates the effect of transitioning from a reference household situation to several comparison household situations (w_k), on income satisfaction, holding income and age constant, while controlling for time-invariant household heterogeneity. Note again that the model is based on conditional maximum likelihood estimators following the blow-up-and-cluster strategy to identify individual thresholds. α^w indicates the scaling component of the replacement rate. Again, if these are negative, the replacement rate decreases with income. If they are non-significant, the replacement rate can be assumed to be income-independent.

In Table 2, all α's for couple households that do not include unemployed partners are negative and significant on the 1% level. The coefficients are similar to the values found by Biewen and Juhasz (2017), e.g., for couples without children as reference. They again, support the finding from the equivalence scales literature that economies of scale are greater at higher income levels. Most importantly, however, the scaling component of the retired single a_{ret} is negative, and is significant on the 1% level. The metric can be interpreted as the benchmark replacement rate, which, according to this finding, decreases with income for single households. The same is true for couple households. The respective α for the households with at least one retiree, aa_{ret} $aa_{ret;un}$, $aa_{ret;emp}$ are all negative, with the dual retiree household having

[11] When adding a quadratic term of log income on the right hand side of the standard model displayed in Fig. 1, the coefficient turns out to be significant.

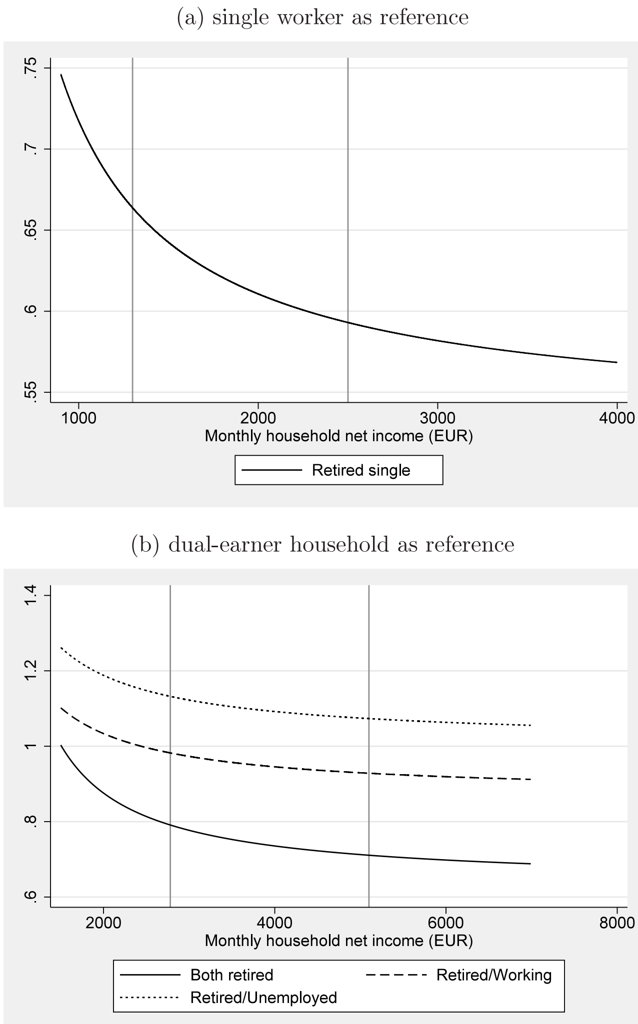

Fig. 5. The replacement rate that maintains income satisfaction.
Note: The estimates are based on the coefficients outlined in Tables 2 and 3. The vertical lines display the interquartile range of income for working individuals aged 50 and older.
Source: German Socio-Economic Panel 1991–2017.

the largest coefficient. They are significant for both aa_{ret} (p=0.002) and $aa_{ret;un}$ (p=0.033), but not $aa_{ret;emp}$ (p=0.171).

ρ^{w} indicates the translation component of the replacement rate, which represents the variable costs of the household transition. In Tables 2 and 3, they are all above 1 except for household situations in which unemployed individuals are involved.

Fig. 5 shows the results of Eq. (12) plotted against household income.[12] A single worker is used as the reference in Panel (a) and a dual-earner household is used as the reference in Panel (b). The gray vertical lines display the 25% and 75% Quartiles of the earning distribution with respect to the sample in employment. On the vertical axis, $r(w_{it})$ represents the replacement rate that maintains income satisfaction based on a household situation in employment.

Among singles, the replacement rate falls, within the interquartile range of earnings, from 67% for monthly household net earnings of around EUR 1300 to about 59% for earnings around EUR 2500. For couple households, the results are different depending on the labor force status of the retiree's partner. If the comparison household consists of two retirees, r falls from 80% for joint earnings of around EUR 2800 to 72% for earnings around EUR 5100. If the retiree's partner is still in employment, r falls from 98% to 92%. If the retiree's partner is unemployed, r falls from 115% to 108%, however, as discussed above, the decline is non-significant.[13]

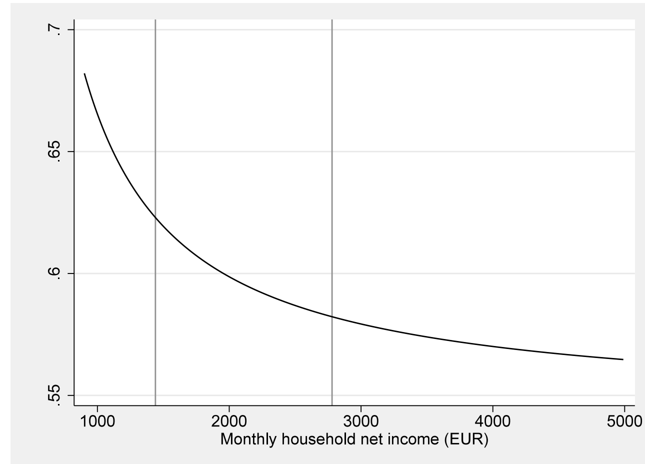

Fig. 6. The replacement rate that maintains income satisfaction; only men; single worker as the reference category.
Note: The estimates are based on the coefficients outlined in Table 2.
Source: German Socio-Economic Panel 1991–2017.

Heterogeneity

In Fig. 6 and Table 4, I show results based on men only, with single working men as the reference category. For men, replacement rate are easier to interpret because current cohorts have less gaps in their earning histories, which is crucial for income from statutory pensions.

The scaling component α decreases to −335 leading to a flatter decline in the benchmark replacement rate than in the main result. The replacement rate now falls from 62% to 58% within the interquartile range of the corresponding earning distribution. However, the decline is not significant as standard errors have increased substantially. As a result, for single men, the replacement rates may be assumed to be income independent.[14]

In the former GDR, earnings were generally lower, and while retirement incomes were also lower, current cohorts from the former GDR still realize higher replacement rates. This is partly because contribution years earned in the former GDR have a different financial value than those earned in West Germany (see Kluth and Gasche, 2015, for a stratified analysis of replacement rates from statutory pensions). In Table 4 and Fig. 7, I show results for individuals that were living in Eastern Germany. For the benchmark replacement rate for singles, the coefficient does not change much, but the results are now less significant (see row 2 in Table 4). For couples, the results are also similar while the standard errors are a bit higher due to the smaller size. Still, the benchmark replacement rate decreases from 90% at a joint income of about EUR 2400 to 80% at a joint income of EUR 4100

[12] I do not plot households with incomes less than the minimum pension income in Germany, which is around EUR 900 for singles and around EUR 1200 for couples. In Germany, when individuals have not contributed enough, their benefits are raised to a minimum pension value. This results naturally in high replacement rates.

[13] Note, that individual replacement rates that maintains income satisfaction through retirement, for persons living in a couple household, cannot be derived from this results. Satisfaction refers to pooled income while the household source income is heterogeneous.

[14] Due to small group sizes, some of the household dummies had to be removed from the specification.

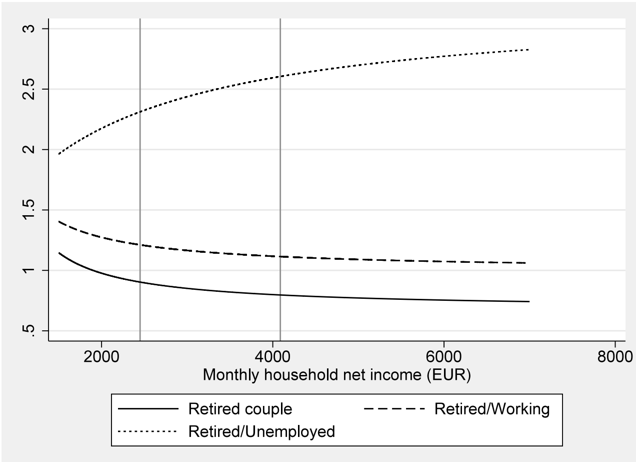

Fig. 7. The replacement rate that maintains income satisfaction; only East Germany (the former GDR); double-earner household as the reference category.
Source: German Socio-Economic Panel 1991–2017.

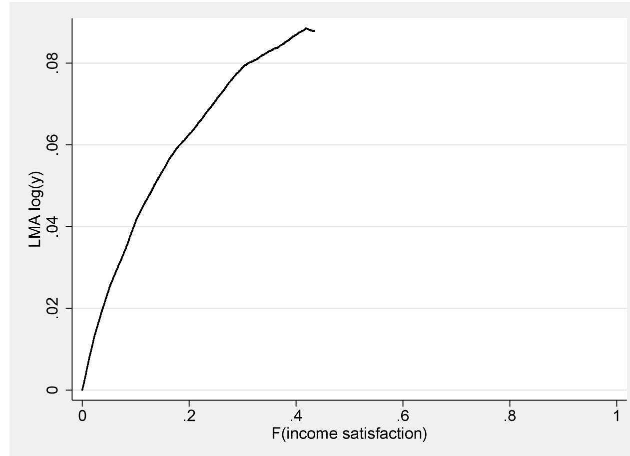

Fig. 8. Line of independence minus absolute concentration curve of household income with respect to income satisfaction.
Source: German Socio-Economic Panel 1991–2017.

Table 4
Scaling component α across applied models.

	Retired single	Retired couple
Main result	−486.8(127.4)***	−871.6(287.0)***
Only men	−335.3(497.0)	(k)
East Germany	−500.8(267.1)*	−836.5(480.6)*
Transformed outcome	−504.5(146.0)***	−788.2(321.3)**
Control for health changes	−463.2(120.2)***	−769.2(290.5)**
Control for changes in homeownership	−486.7(127.4)***	−792.2(298.3)*
Age threshold at 60	−227.9(47.8)***	84.6(169.2.0)
Age interval 60–70	−186.2(84.1)**	−151.8(354.0)

Note: clustered standard errors in parenthesis * represent significance levels; (k) did not converge.

(see Fig. 7). It also shows that rising replacement rates are possible in this setting. Here, for the household type where the retiree's partner is unemployed (see dotted line).

Robustness

How the model uses information from self-reported income evaluations is discussed in Sections "Methodology" and "Dependent variable". Still, I make the inevitable assumption that a cardinal notion of individual utility is behind all of the subjective indications represented by the data (Dudel et al., 2016). It is therefore reassuring if models with weaker assumptions and/or different settings come to the same conclusions. In the following section, I test how sensitive the benchmark replacement rates are to transformations of the dependent variable, different identification strategies, different sample selections, and the inclusion of other time-varying factors.

First, I transform the dependent variable into three categories, with 0–4 indicating *unsatisfied*, 5–7 indicating *neither unsatisfied nor satisfied* and 8–10 indicating *satisfied*. The results are shown in Table 4. For singles, α increases slightly while standard remain relatively small. For couples, α is more or less unchanged but standard errors are larger than in the main model.

Using the procedure suggested by Schröder and Yitzhaki (2017), I check whether the key coefficients of the model are sensitive to further transformations. The authors use what they call a Line of Independence Minus Absolute concentration (LMA) to check whether sign of coefficients flip when using different assumptions about cardinality. Fig. 8 shows the LMA for household income from Eq. (1) plotted against the density of income satisfaction. It shows that there are no breaks into negative levels at any point in the distribution. Thus, there is no possibility that a transformation would change the sign of the coefficients. Age is also tested with the same conclusion.

It is known from the literature that control variables have little impact on the subjective equivalence scales (Schwarze, 2003; Biewen and Juhasz, 2017). Moreover, with the fixed-effects design, unobserved time-invarying factors are accounted for. Still, there are observable factors that might change around retirement, and it is important to check whether including them in the model changes the results.

Health is known to change in retirement. While the literature on whether the change is positive or negative is inconclusive, health is an important factor. Thus, I add health to Eq. (1). I use health satisfaction, which is dichotomized with satisfaction with health being lower than five indicating *bad health*. The coefficient of bad health on income satisfaction is −0.5966, and is thus quite a large effect. However, it does not substantially change the main picture as none of the resulting ρ and α are significantly different from those of the main model (see Table 4).

Another important factor that might change in retirement is homeownership. Retirees who own a home have fewer fixed costs than those who rent an apartment or a house. Therefore they might be more satisfied with their income than a tenant because their expenditures are lower. The fixed-effects approach captures the effect when it is constant over the observation period. Some households may change their status when they retire, by, for example, selling their house and becoming a tenant of a smaller apartment. Thus, I add an indicator variable for tenant or owner to the main model. The effect of this variable on income satisfaction is not distinguishable from zero and it also results in estimates similar to those in the main model (see Table 4).

In Germany, the average income in the final years before retirement is usually lower because many pre-retirees reduce their working hours. As a result replacement rates tend to be higher when this time span is used as the reference.

In another robustness test, I examine to what extent results are sensitive to introducing a narrower age interval. For the main results, I examined adults 50 years or older in order to allow for a reasonable range of observations. In the replacement rate literature is has been argued that replacement rates should refer to the very last years of the respondents' careers. Replacement rates that are realized in Germany and that refer to the final years are usually higher because many pre-retirees reduce their working hours prior to retirement. That makes the share that retirement income needs to replace smaller. When I trim the sample to people aged 60 or older and redo the analysis for singles, I find that replacement rates that maintain income satisfaction are shifted upwards (see Fig. 9). They now fall from 79% to 74% within the

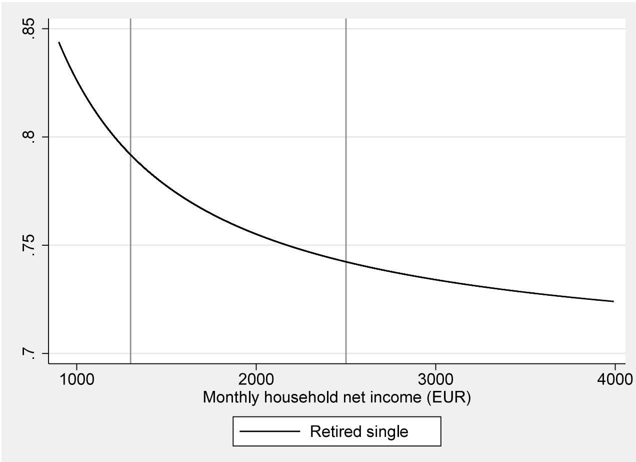

Fig. 9. The replacement rate that maintains income satisfaction; single earner as reference; age threshold at 60.
Source: German Socio-Economic Panel 1991–2017.

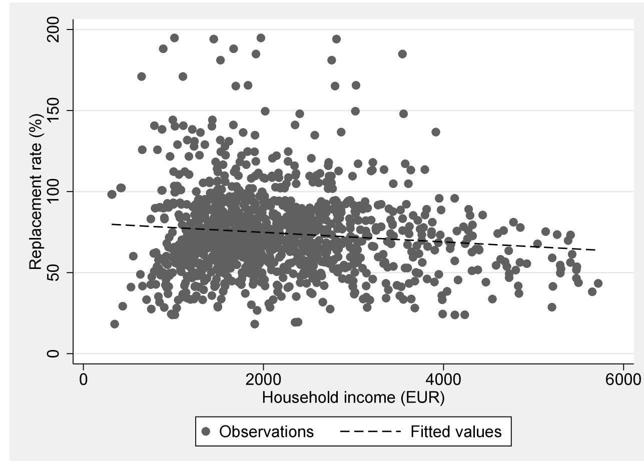

Fig. 10. Empirical replacement rates for satisfied retirees (only singles).
Note: Each dot represents one household. Replacement rates are calculated by dividing the first household income after retirement by the final working income before retiring.
Source: German Socio-Economic Panel 1991–2017.

interquartile range. For couple households, using this threshold results in a positive but very small α which is not significant on conventional levels.

In the main sample, individuals are aged up to 90 years (see Appendix B). At this age, satisfaction levels might be mostly dominantly determined by health. Thus, in a final robustness check, I use a narrow age range from 60 to 70 years (see bottom row of Table 4). Results for singles do not differ much from the test based on the age threshold at 60 years. Results for couples are, again, negative but insignificant.

The identification strategy outlined in Section "Methodology" nicely supports the idea of maintaining the living standards individuals had during their working years (Dudel et al., 2016), but the estimation procedure is complicated and the implementation is tricky (code is available upon request). Moreover, the computational effort is substantial. To address concerns that the main results are not a statistical artifact from the methodology, I demonstrate an alternative way to identify benchmark replacement rates.

Coming back to the initial question what a good choice of replacement rate in the retirement plan would be, we could simply calculate the empirical replacement rates; i.e., the rates that retired households realized. However, using administrative data, there is no obvious way of assessing whether the replacement rate they realized were high enough. Thus, I use the household income from SOEP, calculate the replacement rates for recent cohorts, and select on retirees who have more than median income satisfaction after retirement. The assumption is that retirees who evaluated their household income with value higher than seven on a scale from zero to 10 are satisfied. It turns out that, plotted against income, the replacement rate also decreases with income, with values similar to those in the main results (For simplicity, I only calculated singles), see Fig. 10.

To sum up the findings of the sensitivity analysis, many modifications of the default strategy lead to similar results. The benchmark replacement rate for singles turns out to be fairly robust. While the exact relationship between the metric and income is not equal across models, it is consistently negative, which indicates that the benchmark replacement rates are higher at lower incomes and vice versa. One notable exception is the sample of men, where income independence could not be rejected. For couple households, the standard errors are higher, leading to mostly less significant results. Still, the scaling component is also always negative (See Table 4).

Goodness-of-fit

In this section, I compare the fit of the main model to more general models with stronger assumptions. In particular, I compare the AIC of the model using the BUC strategy and a binary fixed-effects estimator implementing an income-dependent identification as outlined in (1) against an income-independent identification estimated with a fixed-effects ordered logit model.

I calculate the respective Akaike Information Criteria from the log likelihood as follows:

$$AIC = -2 \times (LL) - 2k \quad (13)$$

where k indicates the number of model parameters which is equal across the two models.

The log likelihood of the main model amounts to 114,036 whereas the base independent model results in 193,969. Thus the AIC of the main model is smaller than the AIC of the more general form, indicating that the fit of the main model is superior.

Discussion

Comparison to existing results

The income-dependent replacement rates calculate above are substantially lower than the rates found by Dudel et al. (2016), in particular, for the more affluent singles. In comparison to Binswanger and Schunk (2012), my results are surprisingly similar to their finding for the Netherlands, (see Section "Related literature").

Comparing my results with realized replacement rates from recent cohorts in Germany should be pursued with care. For empirical replacement rates, when they are available based on register data, the levels strongly depend on the reference time span applied. Using a similar time span to that in this paper, Kluth and Gasche (2015) found replacement rates of around 50 percent. Across income classes, they only provided, what they called, life-cycle replacement rates across income classes, which implies, that the reference period was the average income from the life-cycle. For West Germany, these replacement rates increased with income from 30 to 45 percent (they only considered individuals). Note however, that the authors did not take into account income from wealth. Still, it is important to note that the realized replacement rates were lower for low earners while the results of this paper suggest that they should be higher for low-earners. The opposite is the case for Eastern Germany, where replacement rates are higher for low earners (Kluth and Gasche, 2015).

Explanation angles

When households transition into retirement their expenses usually change (e.g. Bernheim et al., 2001; Schwerdt, 2005). In particular, their work-related costs, such as their expenditures on commuting, business clothes, accommodation or travel tend to decrease or drop after retirement. Also savings, when regarded as a form of costs, drop or at least decrease upon retirement. Conversely, because households have fewer time constraints, other types of expenses, particularly for leisure, typically increase. Finally, households' age-related costs, such as their health expenditures tend to increase, because they are, on average, older in retirement than when working. A priori, it is unclear whether those changes in expenses are proportionally larger or smaller across income classes. However, if the decrease in retirement costs is, in sum, larger for more affluent households, they would need less of their income to achieve their desired utility level than poorer households would. There is some evidence, from the US, that people with low wealth accumulations experience higher expenditures cuts after retirement (due to poor health), whereas for the richer part expenditures increase (Moran et al., 2021). For Germany, Schwerdt (2005) finds that people with high replacement ratios have consumption increases after retirement whereas for low replacement ratios consumption drops up to 30%.

Recent literature has suggested that as people spend more time in retirement, they perform activities such as shopping and meal preparation more efficiently (Aguiar and Hurst, 2005; Luengo-Prado and Sevilla, 2013). Therefore, they need a smaller amount net income to achieve the same utility level that they did while working. It may be the case that those efficiency gains are more pronounced among the more affluent households. Another potential explanation for my findings, albeit only for couple households, is that retired couples benefit from economies of scale. From the equivalence scale literature we know that economies of scales are greater among richer households. It is possible that more affluent couples are able to increase their economies of scale after retirement, and are thus able to achieve higher utility with less income.

Policy implications

The results of this study show that, after retirement, lower-income households have to make an extra effort to maintain themselves financially at the welfare level they had before retirement (which might have been low in the first place; see limitations section below). Extra effort implies, that while they are working, these households have to consume less of their net income to accumulate wealth, as they have replace proportionally more of their earnings than higher earners do. However, given that they usually have a lower capacity to save than high earners, because they have less wealth and higher fixed costs, they might not be able to do so on their own (Crossley and O'Dea, 2010).

Many pension systems have a tax-financed redistributive component in which low earners are supported by the high earners (e.g., Sweden, Denmark). Others have a large progressive component where larger contributions result in relatively smaller benefits.

In Germany, the constant benefit-to-contribution rates are historically deeply rooted in the Bismarckian system. Following the "Äquivalenzprinzip", pension benefits are to a large degree determined by earning contributions, with high contributions being rewarded with high pension benefits (until a certain income level, beyond which the system is no longer mandatory). However, parts of the system have already been changed under the current government.

Conclusion

In this paper, I addressed the question of what a good choice of a replacement rate would be when making retirement saving plans. I found that the replacement rates that maintain income satisfaction through retirement are higher for households with lower incomes and lower for household with higher incomes. That was shown to be the case for both single and couple households, whereby the results for singles are more precise and consistent across a wide range of changes in the applied approach. While I offered some potential explanations, I leave further investigations of these findings to future research.

It is important to stress the relative dimension of this analysis. Maintaining income satisfaction does not rule out the possibility that a household has difficulties to making ends meet. Therefore, poverty thresholds should always be taken into account additionally (e.g. Love et al., 2008). Furthermore, while income from wealth was considered, households may hold wealth that has not been reported. Similarly, potential inheritances could increase the households' retirement income were not taken into account.

While income-dependent benchmark replacement rates as found in this paper might be more accurate, a single benchmark as in Dudel et al. (2016) or Dudel and Schmied (2019) is easier to remember by savers and a better rule of thumb.

Research on benchmark replacement rates is still scarce, which is surprising given their practical relevance. This approach could be applied to other countries, where datasets include income satisfaction. Meanwhile, measuring realized, desired, and welfare maintaining replacement rates for the very same households is a potential avenue for future research.

Declaration of competing interest

None.

Acknowledgments

I am grateful to Rob Alessie for motivating me to work on this topic; Timm Bönke, Christian Dudel, Carsten Schröder and Martin Werding, for helpful comments, and Martin Biewen for providing code to Biewen and Juhasz (2017). I also thank two anonymous referees and the seminar participants of the BeNA Summer Workshop 2020.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), grant DU 1566/1-1.

Appendix A: Fit of functional form

See Fig. 11.

Appendix B: Summary statistics

See Table 5.

Appendix C: Implementation of the blow-up-and cluster strategy

I assume that the reported levels of income satisfaction are generated by an ordered logit model, such as:

$$P(u_{it} = k|z_{it}, w_{it}) = \Lambda(\tau_k - \lambda logQ(b,a) - z_{it}\beta - \alpha_i) - \Lambda(\tau_{k-1} - \lambda logQ(b,a) - z_{it}\beta - \alpha_i) \quad (C.1)$$

where $Q(b,a)$ represents the replacement rate that sets u before and after the retirement transition equal; τ_k represents a threshold for individual i who makes a certain subjective evaluation of his/her income.

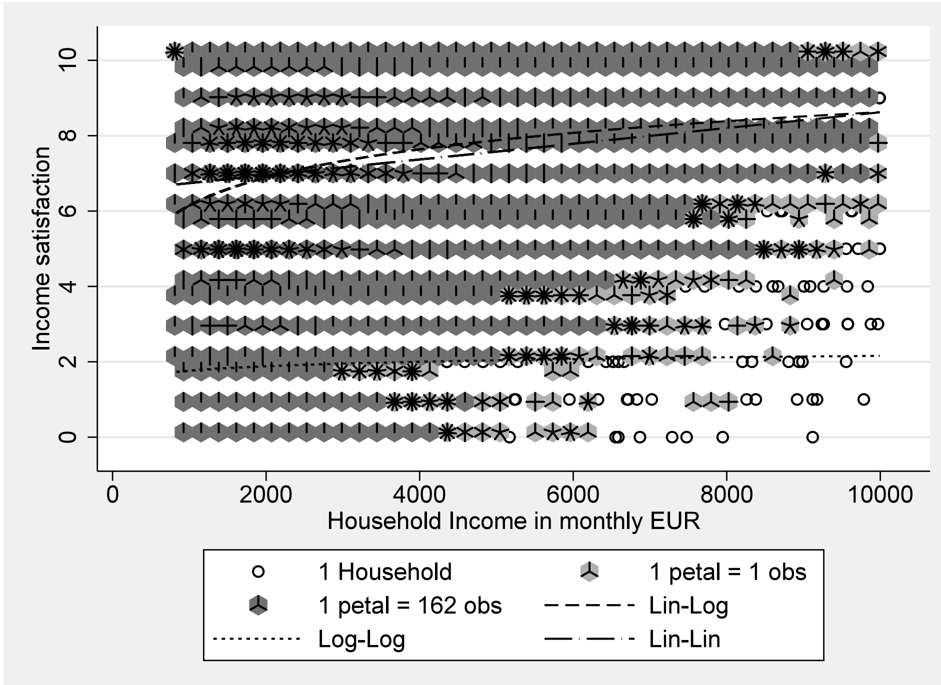

Fig. 11. Sunflower graph to show fit of functional forms.
Source: German Socio-Economic Panel 1991–2017.

Table 5
Sample summary statistics.

Percent Female (in %)	50,5
Year of birth (min)	1923
Year of birth (max)	1957
Share singles at age 65 (%)	12.5
Share singles at age 80 (%)	36.0
Age (min)	50
Age (max)	92
Retirement year (mean)	63.4
Retirement year (min)	60
Retirement year (max)	69

In the BUC method, u is dichotomized in $k-1$ ways, where $k=11$ as income satisfaction is scaled 0–10 in the SOEP. Copies for all observations are generated using different cut-off points. The CML is applied to any dichotomization. The BUC method repeats this procedure for all possible combinations, while restricting the estimates to be equal across all dichotomizations. (see Dudel et al., 2016 for more details).

References

Aguiar, M., Hurst, E., 2005. Consumption versus expenditure. J. Polit. Econ. 113, 919–948.
Baetschmann, G., Staub, K.E., Winkelmann, R., 2015. Consistent estimation of the fixed effects ordered logit model. J. R. Stat. Soc.: Ser. A (Statistics in Society) 178 (3), 685–703.
Bernheim, B.D., Skinner, J., Weinberg, S., 2001. What accounts for the variation in retirement wealth among U.S. households? Amer. Econ. Rev. 91, 832–857.
Bertrand, M., Mullainathan, S., 2001. Do people mean what they say? Implications for subjective survey data. Amer. Econ. Rev. 91 (2), 67–72.
Biewen, M., Juhasz, A., 2017. Direct estimation of equivalence scales and more evidence on independence of base. Oxf. Bull. Econ. Stat. 79, 875–905.
Binswanger, J., Schunk, D., 2012. What is an adequate standard of living during retirement? J. Pension Econ. Finance 11, 203–222.
Blackorby, C., Donaldson, D., 1993. Adult-equivalence scales and the economic implementation of interpersonal comparisons of well being. Soc. Choice Welf. 10, 335–361.
Ferrer-i Carbonell, A., Frijters, P., 2004. How important is methodology for the estimates of the determinants of happiness? Econ. J. 114 (497), 641–659.
Chamberlain, G., 1979. Analysis of Covariance with Qualitative Data. National Bureau of Economic Research.
Chiappori, P.-A., 2016. Equivalence versus indifference scales. Econ. J. 126 (592), 523–545.
Crawford, R., O'Dea, C., 2012. The Adequacy of Wealth Among Those Approaching Retirement. IFS Report R72, Institute for Fiscal Studies, London.
Crossley, T.F., O'Dea, C., 2010. The wealth and saving of UK families on the eve of the crisis. IFS Reports R71, IFS reports, Institut for Fiscal Studies, London.
Deaton, A.S., Muellbauer, J., 1980. An almost ideal demand system. Am. Econ. Rev. 70 (3), 312–326.
Deaton, A., Muellbauer, J., 1986. On measuring child costs: With application to poor countries. J. Polit. Econ. 94, 720–744.
Donaldson, D., Pendakur, K., 2006. The identification of fixed costs from consumer behaviour. J. Bus. Econom. Statist. 24, 255–265.
Dudel, C., Ott, N., Werding, M., 2016. Maintaining one's living standard at old age: What does that mean? Empir. Econ. 51, 1261–1279.
Dudel, C., Schmied, J., 2019. Pension Adequacy Standards: an Empirical Estimation Strategy and Results for the United States and Germany. MPIDR Working Paper 3, Max Planck Institute for Demographic Research.
Geyer, J., Steiner, V., 2014. Future public pensions and changing employment patterns across birth cohorts. J. Pension Econ. Finance 13 (2), 172–209.
Grabka, M., 2020. SOEP core v36: Codebook for PEQIVAL. SOEP Survey Papers ; 922: Series D - Variable Description and Coding.
Kahneman, D., Krueger, A.B., 2006. Developments in the measurement of subjective well-being. J. Econ. Perspect. 20 (1), 3–24.
Kluth, S., Gasche, M., 2015. Ersatzraten in der Gesetzlichen Rentenversicherung/Replacement Rates in the German Statutory Pension System. Jahrbücher für Nationalökonomie und Statistik 235 (6), 553–583.
Knoef, M., Been, J., Alessie, R., Caminada, K., Goudswaard, K., Kalwij, A., 2016. Measuring retirement savings adequacy: Developing a multi-pillar approach in the Netherlands. J. Pension Econ. Finance 15, 55–89.
Koulovatianos, C., Schröder, C., Schmidt, U., 2005. On the income dependence of equivalence scales. J. Public Econ. 89 (5–6), 967–996.
Krueger, A.B., Schkade, D.A., 2008. The reliability of subjective well-being measures. J. Public Econ. 92 (8–9), 1833–1845.
Layard, R., Mayraz, G., Nickell, S., 2008. The marginal utility of income. J. Public Econ. 92 (8–9), 1846–1857.
Lewbel, A., 1989a. Household equivalence scales and welfare comparisons. J. Public Econ. 39, 377–391.
Lewbel, A., 1989b. Nesting the AIDS and translog demand systems. Internat. Econom. Rev. 30 (2), 349–356.
Love, D.A., Smith, P.A., McNair, L.C., 2008. A new look at the wealth adequacy of older U.S. households. Rev. Income Wealth 54, 616–642.
Luengo-Prado, M.J., Sevilla, A., 2013. Time to cook: Expenditure at retirement in Spain. Econom. J. 123, 764–789.

Moran, P., O'Connell, M., O'Dea, C., Parodi, F., Submitter, M.R., et al., 2021. Heterogeneity in Household Spending and Well-being around Retirement. Working Paper, University of Michigan Retirement and Disability Research Center (MRDRC).

OECD, 2015. Pensions at a glance 2015. Available online at http://www.oecd-ilibrary.org/social-issues-migration-health/pensions-at-a-glance-2015_pension_glance-2015-en.

Palmer, B.A., 1989. Tax reform and retirement income replacement ratios. J. Risk Insurance 56 (4), 702–725.

Pradhan, M., Ravallion, M., 2000. Measuring poverty using qualitative perceptions of consumption adequacy. Rev. Econ. Stat. 82 (3), 462–471.

Pudney, S., 2011. Perception and retrospection: The dynamic consistency of responses to survey questions on wellbeing. J. Public Econ. 95 (3–4), 300–310.

Scholz, J., Seshadri, A., Khitatrakun, S., 2006. Are Americans saving 'optimally' for retirement? J. Polit. Econ. 114, 607–643.

Schröder, C., Yitzhaki, S., 2017. Revisiting the evidence for cardinal treatment of ordinal variables. Eur. Econ. Rev. 92, 337–358.

Schulz, J., Carrin, G., 1972. The role of savings and pension systems in maintaining living standards in retirement. J. Hum. Resour. 7, 343–365.

Schwarze, J., 2003. Using panel data on income satisfaction to estimate equivalence scale elasticity. Rev. Income Wealth 49, 359–372.

Schwerdt, G., 2005. Why does consumption fall at retirement? Evidence from Germany. Econom. Lett. 89 (3), 300–305.

Skinner, J., 2007. Are you sure you're saving enough for retirement? J. Econ. Perspect. 21 (3), 59–80.

Smith, J.P., 2003. Trends and projections in income replacement during retirement. J. Labor Econ. 21 (4), 755–781.

UK Pension Commission, 2004. Pensions: Challenges and choices; The first report of the pensions commission.

van Praag, B.M.S., Kapteyn, A., 1973. Further evidence on the individual welfare function of income. Eur. Econ. Rev. 4, 32–62.